DYNAMICAL SYSTEMS AND CONTROL THEORY INSPIRED BY MOLECULAR BIOLOGY

FA9550-08-1-0053

Eduardo Sontag, Department of Mathematics, Rutgers University

Abstract:

This project aims to develop new concepts, theory, and algorithms for control and signal processing using ideas inspired by molecular systems biology. Cell biology provides a wide repertoire of systems that are strongly fault-tolerant, nonlinear, feedback-rich, and truly hybrid, while making effective use of highly heterogeneous sensing and actuation channels. Advances in genomic/proteomics and molecular systems biology research are continually adding detailed knowledge of such systems' architecture and operation, thus offering, in principle, a powerful source of inspiration for innovative solutions to problems of control and communication, sensor and actuator design, and systems integration. Often, problems in systems biology may superficially resemble standard problems in control theory but, on closer inspection, they differ in fundamental ways. These differences are challenging and worth exploring, and lead to and interesting and highly rewarding new directions of research.

Summary of Findings:

Systems that are "close" to monotone.:

Monotone systems (which arose from the study of gene regulation and other biological systems) have proved a powerful tool for analyzing many nonlinear systems. However, monotonicity is a restrictive property vis a vis general biological as well as engineering systems. During the past few years, the PI and his collaborators and students have developed an approach (monotone I/O theory) that allows one to exploit the monotonicity of components even when the overall system is not monotone. In this context, a main line of work in this project is that of understanding the dynamics of systems that are not necessarily monotone but which, in some manner, are "close" to being so. In monotone systems, every net feedback loop is positive. On the other hand, negative feedback loops are important features of many systems, since they are required for adaptation and precision. The paper [19] used geometric singular perturbation theory to show that, provided that these negative loops act at a comparatively fast time scale, the main dynamical property of (strongly) monotone systems, convergence to steady states, is still valid. An application was worked out to a double-phosphorylation "futile cycle" motif which plays a central role in eukaryotic cell signaling. This work was complemented by [18], which studied the number of positive steady states in biomolecular reactions consisting of activation/deactivation futile cycles, such as those arising from phosphorylations and dephosphorylations at each level of a MAPK cascade. It was shown that for some parameter ranges, there are at least $n+1$ (if n is even) or n (if n is odd) steady states, there never are more than $2n-1$ steady states, that for parameters near the standard Michaelis-Menten quasi-steady state conditions, there are at most $n+1$ steady states and that for parameters far from the standard Michaelis-Menten quasi-steady state conditions, there is at most one steady state.

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
20 FEB 2011		**02-01-2008 to 30-11-2010**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Dynamical Systems And Control Theory Inspired By Molecular Biology	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Rutgers, The State University of New Jersey,New Brunswick,NJ,08901**	8. PERFORMING ORGANIZATION REPORT NUMBER **; AFRL-OSR-VA-TR-11-0034**
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S) **AFRL-OSR-VA-TR-11-0034**

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

This project aims to develop new concepts, theory, and algorithms for control and signal processing using ideas inspired by molecular systems biology. Cell biology provides a wide repertoire of systems that are strongly fault-tolerant, nonlinear, feedback-rich, and truly hybrid, while making effective use of highly heterogeneous sensing and actuation channels. Advances in genomic/proteomics and molecular systems biology research are continually adding detailed knowledge of such systems' architecture and operation, thus offering, in principle, a powerful source of inspiration for innovative solutions to problems of control and communication, sensor and actuator design, and systems integration.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
unclassified	**unclassified**	**unclassified**	**Same as Report (SAR)**	**9**	

Cooperative systems are the most common particular type of monotone system. Their attractors are particularly simple; for example, a nontrivial periodic orbit cannot be an attractor. The paper [1] provided characterizations of attractors for the wider class of "coherent" systems, defined by the property that all directed feedback loops are positive, generalizing the above and other results. The class of coherent systems serves to model many biological systems. Several new results for cooperative systems were obtained in the process as well.

Identifiability of nonlinear systems.:

In molecular biology, typically only very simple input signals (such as steps or pulses) can be used to perturb systems, which brings up the issue of identifiability when only using such inputs. Interestingly, this most natural question seems to have attracted very little attention in control theory. In [16], we dealt with the following question: what classes of input signals are sufficient in order to completely identify the input/output behavior of generic bilinear systems? (Bilinear systems are commonly used models in control applications, as well as in biology.) The main results are that (1) step inputs are not sufficient, nor are single pulses, but (2) the family of all pulses (of a fixed amplitude but varying widths) do suffice for identification. This project is being actively pursued, with many questions left open, including the possibility of proving similar results for other classes of inputs, as well as generalizing to other classes of (not necessarily bilinear) systems.

Feedforward Circuits, Adaptation:

"Feedforward" circuits are ubiquitous in biology. For example, (incoherent) feedforward circuits are over-represented in many gene transcription networks (compared to other "motifs" involving three nodes), as well as in control mechanisms in mammalian cells and at the interface of genetic and metabolic networks. A large number of papers have been devoted to the signal-processing capabilities of the feedforward motif. In biology, feedforward circuits have been also often proposed for adaptation (disturbance rejection) of constant signals. Indeed, a popular such circuit, implemented by a particular chemical reaction scheme (the "sniffer") has been proposed as the paradigm for perfect adaptation, and several authors have identified roles for such circuits in the adaptation by the chemotaxis pathway of *Dictyostelium*, in protein expression regulation by microRNA-mediated loops, and in explaining the robust behavior of *E.coli* carbohydrate uptake via the carbohydrate phosphotransferase system. Although lack of feedback might make such circuits lack robustness, they typically enjoy, for appropriate parameter choices and algebraic forms of reactions, an almost-disturbance property.

In [13], we started a careful analysis of feedforward circuits for adaptation. A global convergence theorem was proved in a general framework, which includes examples from the literature as particular cases. An interesting observation was that, in contrast to integral-feedback based adaptation, these circuits do not adapt to pulse signals, displaying a memory phenomenon.

The paper [2] developed experimental synthetic-biology transcriptional and post-transcriptional incoherent feedforward circuits to test the adaptability properties of such feedforward circuits. In mammalian cells, three circuits were designed. In one, a transcriptional repressor is co-expressed together with the unit's output and downregulates this output via DNA binding. In two other circuits, a negative regulator is a microRNA that targets output mRNA via RNA interference. We compared the incoherent feed-forward motifs with a negative autoregulation motif and analyzed

their differences with respect to adaptivity in copy number and robustness to noise. Experiments show that the transcriptional and post-transcriptional incoherent feedforward motif show superior adaptivity to copy number as compared to the negative autoregulation and the noise levels depend on the type of regulation yet again showing more robust behavior from the post-transcriptional implementations. The paper [12] also dealt largely with feedforward systems, as well as related constructions. The starting observation was that many cellular sensory systems display fold-change detection (FCD): a response whose entire shape, including amplitude and duration, depends only on fold-changes in input, and not on absolute changes. Thus, a step change in input from, say, level 1 to 2, gives precisely the same dynamical output as a step from level 2 to 4, since the steps have the same fold-change. We ask what is the benefit of FCD, and show that FCD is necessary and sufficient for sensory search to be independent of multiplying the input-field by a scalar. Thus the FCD search pattern depends only on the spatial profile of the input, and not on its amplitude. Such scalar symmetry occurs in a wide range of sensory inputs, such as source strength multiplying diffusing/convecting chemical fields sensed in chemotaxis, ambient light multiplying the contrast field in vision, and protein concentrations multiplying the output in cellular signaling-systems. Mathematically, we provided an effective characterization of the FCD property via symmetries in state-space, and showed that FCD implies two features found across sensory systems, exact adaptation and Weber's law, but that these two properties are not sufficient for FCD. We also analyzed a wide class of mechanisms that have FCD, including certain non-linear feedback and feedforward loops. We found that bacterial chemotaxis displays feedback within the present class, and hence is expected to show FCD. This can explain experiments in which chemotaxis searches are insensitive to attractant source levels. This study thus suggested a connection between properties of biological sensory systems and scalar symmetry stemming from physical properties of their input-fields. A far-reaching generalization to arbitary symmetries, whose proof relies upon deep results in nonlinear control theory, gave a general characterization in terms of solutions of certain partial differential equations. A paper reporting on these results is currently under review [11].

Networks and contraction:

It is often very useful to break down the analysis and design of large-scale systems into two independent steps: At a "global" level, properties of a network or interconnection graph are imposed so as to guarantee a desired behavior for the full interconnected system. In this analysis, subsystems may be characterized as "black boxes" with assumed input-output characteristics, but detailed knowledge of their internal structure is not required. At a "local" level of analysis, one imposes constraints on the structure and behavior of individual subsystems (components), so as to fit the requirements of the global approach. These requirements are verified independently of the overall network structure. This "multi-scale" or hierarchical methodology is robust in so far as a large degree of uncertainty can be tolerated in the components, only constrained by meeting appropriate behavioral requirements. There are many examples of such approaches in control theory, including among others: the use of small-gain theorems to guarantee stability of a negative feedback loop provided that the components are individually stable (qualitative property of components) and the overall loop H_∞ gain is less than one, as well as nonlinear generalizations based on input to state stability; input/output monotone systems theory, in which input-output characteristics are the only required "quantitative" data; and the use of passivity-based tools.

In [9] (a journal paper is under preparation as well), we present yet another example of the gen-

eral principle, this time in the framework of contractive systems. We show that contractivity of individual components coupled with contractivity of an appropriate interconnection matrix (which uses only norm and matrix-measure estimates, but no precise knowledge, of components) guarantees contractivity of the overall system. No assumptions are made on the networks. Directed networks, self-loops, and multiple regulatory interactions are allowed. This paper followed-up on our work [8], which gave conditions for transcriptional systems to be globally entrained to external periodic inputs. It was shown there that certain systems driven by external periodic signals have the property that all solutions converge to a fixed limit cycle. General results were proved, and the properties were verified in the specific case of some models of transcriptional systems.

Robustness of phenotype to large parameter variations.:

A common quantification of robustness in biology, with some analogies but technically different to the use of the term in control theory, is in terms of the volume of the space of parameters with are "admissible" in the sense of providing the correct "phenotype" (structure and behavior). In our papers [3, 4] and ongoing work, we postulate that not only the volume of the space but also its topology and geometry contain information on essential aspects of the network, including feasible pathways, and the switching between two parallel pathways or distinct/disconnected active regions of parameters. A general method was presented in these papers to characterize the space of admissible parameters, by writing it as a semi-algebraic set, and then theoretically analyzing its topology and geometry, as well as volume. A stochastic parameter-drift method was introduced for characterization of the shape of parameter space. These methods provide a more objective and complete measure of the robustness of a developmental module. As an illustration, the fruit fly segment polarity gene network was analyzed in detail.

Partially observed stochastic systems.:

Hidden Markov Models (finite partially observed stochastic systems) are of great interest in pattern recognition and classfication applications. Motivated by the question of identifying transcription binding sites for the p53 protein, we developed a new computational method ("p53HMM") that uses as prior information the existence of symmetries in the patterns to be recognized. (In the case of the p53 transcription factor, arising from its tetramer form when binding to DNA.) The p53HMM method was developed in [7] and incorporates a novel "Correlated Baum Welch" training algorithm that provides increased predictive power by exploiting the redundancy of information found in the repeated, palindromic p53-binding motif. The predictive accuracy of these new models was compared against other predictive models, including position specific score matrices. (The p53 protein regulates the transcription of many different genes in response to a wide variety of stress signals. Following DNA damage, p53 regulates key processes, including DNA repair, cell-cycle arrest, senescence and apoptosis, in order to suppress cancer; we published a survey of p53 transcription factor binding sites in Nature Reviews Molecular Cell Biology, 2008.)

Synchronization:

The analysis of synchronization phenomena in networks has become an important topic in systems and control theory, motivated by diverse applications in physics, biology, and engineering. Both the individual dynamics of the components and the network structure play an important role in determining conditions leading to synchronization. In [10], and motivated by such questions, we

started the study of synchronization conditions for networks of nonlinear systems based on passivity notions. The components of the network (referred to as "compartments" in the paper) are made up of an identical interconnection of subsystems, each represented as an operator in an extended L_2 space and referred to as a "species". These compartments are, in turn, coupled through a diffusion-like term among the respective species. The motivation was the analysis of cellular networks where signaling occurs both internally, through interactions of species, and externally, through inter-cellular signaling, The synchronization conditions are provided by combining the input-output properties of the subsystems with information about the structure of network. This work was based on our previous work with Arcak that studied stability properties of individual compartments, rather than synchronization of compartments. That work was based on an appropriate passivity property for each species, and introduced a "dissipativity matrix" that incorporates information about the passivity of the subsystems, the interconnection structure of the species, and the signs of the interconnection terms. To determine the stability of the network, one checks the diagonal stability of the dissipativity matrix, similarly to classical work on large-scale systems by Vidyasagar and others. In the special case of a cyclic interconnection structure with negative feedback, the diagonal stability test encompasses the classical "secant criterion" from mathematical biology. The paper [10] also explored results for state-space models, as well as biochemical applications, and the theory was illustrated through the derivation of synchronization conditions for networks of genetic oscillators.

Modularity and interconnections.:

Modularity plays a fundamental role in the prediction of the behavior of a system from the behavior of its components, guaranteeing that the properties of individual components do not change upon interconnection. Just as electrical, hydraulic, and other physical systems often do not display modularity, nor do many biochemical systems, and specifically, genetic networks. Our previous work (Nature Molecular Systems Biology, 2008) started the study of the effect of interconnections on the input/output dynamic characteristics of transcriptional components, focusing on a property, which we called "retroactivity," that plays a role analogous to non-zero output impedance in electrical systems. In transcriptional networks, retroactivity is large when the amount of transcription factor is comparable to, or smaller than, the amount of promoter binding sites, or when the affinity of such binding sites is high. In order to attenuate the effect of retroactivity, we proposed a feedback mechanism inspired by the design of amplifiers in electronics. We introduced, in particular, a mechanism based on a phosphorylation/dephosphorylation cycle. This mechanism enjoys a remarkable insulation property, due to the fast time scales of the phosphorylation and dephosphorylation reactions. Such a mechanism, when viewed as a signal transduction system, has thus an inherent capacity to provide insulation and hence to increase the modularity of the system in which it is placed. Recent work includes an expository paper [5] (see also [6]) framing modularity questions in the context of constructs in synthetic biology, as well as an experimental paper under review for a biology journal. In [14, 15], we remarked that reverse-engineering techniques in systems biology tend to rely upon data on steady-state (or dynamic) perturbations –obtained from siRNA, gene knock-down or overexpression, kinase and phosphatase inhibitors, or other interventions– in order to understand the interactions between different "modules" in a network. We reviewed one such popular such technique, introduced by the author and collaborators (modular response analysis), and focused on why conclusions drawn from its use –including papers published in the biological literature– may be misleading due to retroactivity effects. A

theoretical result characterizing stoichiometric-induced steady-state retroactivity effects was also given for a class of biochemical networks.

Exact calculation of moments for certain stochastic kinetics:

We have recently started research into stochastic aspects in systems biology. Deterministic models represent aggregate or average behavior, and are particularly accurate in classical chemistry, where the numbers of molecules are very large (and expressed in multiples of Avogadro's number $\approx 6 \times 10^{23}$): by the law of large numbers, the mean behavior is a good description of the system. The main advantage of deterministic models is that they are comparatively easier to study than probabilistic ones. However, deterministic models may be inadequate when probabilistic effects are nontrivial. We have already achieved progress along one line of research. Recent work by other authors had shown that, for any complex-balanced reaction network (including any weakly reversible deficiency zero network), the corresponding Chemical Master Equation (CME) has a steady-state solution of a product of Poisson form. (Complex-balanced networks, studied by Feinberg, Horn, and Jackson, are the focus of much of our research into chemical reaction networks, and the above result turns out to be closely related to the "nonlinear traffic equations" in queuing theory.) When there are no stoichiometric constraints, the variables corresponding to numbers of each species are IID Poisson distributed, and all joint moments can be easily computed in closed form. If there are conservation laws, however, the distribution becomes that of IID Poisson variables subject to conditioning by linear equations, and the computation of moments is highly nontrivial. In [17], we introduced a new approach to the fast computation of such moments. This approach is based on considering the un-normalized probability generating function, and observing that it can be expressed as an exponential polynomial. Thus one may express it as a multi-contour integral, and it follows that Wilf-Zeilberger theory of hypergeometric identities can be used to derive multivariable linear recurrences (with polynomial coefficients in each of its arguments) which allow one to recursively evaluate statistics. A Maple package ("MVPoisson") was developed and made freely available on the web to perform such computations. The use of this package provides several orders of magnitude improvements over the stochastic simulation algorithm (SSA).

Acknowledgment/Disclaimer: This work was sponsored (in part) by the Air Force Office of Scientific Research, USAF, under grant/contract number FA9550-08. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the Air Force Office of Scientific Research or the U.S. Government.

Publications

[1] D. Angeli, M.W. Hirsch, and E.D. Sontag. Attractors in coherent systems of differential equations. *J. of Differential Equations*, 246:3058–3076, 2009.

[2] L. Bleris, Z. Xie, D. Glass, A. Adadey, E.D. Sontag, and Y. Benenson. Synthetic incoherent feed-forward circuits show adaptation to the amount of their genetic template. 2011. Submitted, under revision.

[3] M. Chaves, A. Sengupta, and E.D. Sontag. Geometry and topology of parameter space: investigating measures of robustness in regulatory networks. *J. of Mathematical Biology*, 59:315–358, 2009.

[4] A. Dayarian, M. Chaves, E.D. Sontag, and A. Sengupta. Shape, size and robustness: Feasible regions in the parameter space of biochemical networks. *PLoS Computational Biology*, 5:e10000256, 2009.

[5] D. Del Vecchio and E.D. Sontag. Engineering principles in bio-molecular systems: From retroactivity to modularity. *European Journal of Control*, 15:389–397, 2009.

[6] D. Del Vecchio and E.D. Sontag. Synthetic biology: A systems engineering perspective. In B. Ingalls and P. Iglesias, editors, *Control Theory in Systems Biology*, pages 101–123. MIT Press, 2009.

[7] T. Riley, X. Yu, E.D. Sontag, and A. Levine. The P53HMM algorithm: using novel profile Hidden Markov Models to detect p53-responsive genes. *BMC Bioinformatics*, 10:111, 2009.

[8] G. Russo, M. di Bernardo, and E.D. Sontag. Global entrainment of transcriptional systems to periodic inputs. *PLoS Computational Biology*, 6:e1000739, 2010.

[9] G. Russo, M. di Bernardo, and E.D. Sontag. Stability of networked systems: a multi-scale approach using contraction. In *Proc. IEEE Conf. Decision and Control, Atlanta, Dec. 2010*, page FrB14.3, 2010.

[10] L. Scardovi, M. Arcak, and E.D. Sontag. Synchronization of interconnected systems with applications to biochemical networks: an input-output approach. *IEEE Transactions Autom. Control*, 55:1367–1379, 2010.

[11] O. Shoval, U. Alon, and E.D. Sontag. Symmetry invariance for adapting biological systems. 2011. submitted.

[12] O. Shoval, L. Goentoro, Y. Hart, A. Mayo, E.D. Sontag, and U. Alon. Fold change detection and scalar symmetry of sensory input fields. *Proc Natl Acad Sci USA*, 107:15995–16000, 2010.

[13] E.D. Sontag. Remarks on feedforward circuits, adaptation, and pulse memory. *IET Systems Biology*, 4:39–51, 2010.

[14] E.D. Sontag. Remarks on structural identification, modularity, and retroactivity. In *Proc. IEEE Conf. Decision and Control, Atlanta, Dec. 2010*, page ThA23.1, 2010.

[15] E.D. Sontag. Modularity, retroactivity, and structural identification. In H. Koeppl, G. Setti, M. di Bernardo, and D. Densmore, editors, *Design and Analysis of Biomolecular Circuits*. Springer-Verlag, 2011.

[16] E.D. Sontag, Y. Wang, and A. Megretski. Input classes for identification of bilinear systems. *IEEE Transactions Autom. Control*, 54:195–207, 2009.

[17] E.D. Sontag and D. Zeilberger. A symbolic computation approach to a problem involving multivariate Poisson distributions. *Advances in Applied Mathematics*, 44:359–377, 2010.

[18] L. Wang and E.D. Sontag. On the number of steady states in a multiple futile cycle. *Journal of Mathematical Biology*, 57:29–52, 2008.

[19] L. Wang and E.D. Sontag. Singularly perturbed monotone systems and an application to double phosphorylation cycles. *J. Nonlinear Science*, 18:527–550, 2008.

Personnel Supported During Duration of Grant: Eduardo Sontag, Professor, Rutgers University; Liming Wang, Graduate Student, Rutgers University; Bhaskar Dasgupta, Associate Professor, U. of Illinois (as sabbatical visitor); Paola Vera-Licona, Postdoc, Rutgers University; Michael de Freitas, Graduate Student, Rutgers University

Honors & Awards Received: Elected Fellow of Society for Industrial and Applied Mathematics (2009); IEEE Control Systems Award (2011)

AFRL Point of Contact: Dr. Fariba Fahroo, Program Manager, Computational Mathematics AFOSR/RSL, 875 North Randolph Street, Suite 325, Room 3112, Arlington, VA 22203